DOCUMENT RESUME

ED 040 820 RE 002 792

AUTHOR	Follman, John; And Others
TITLE	Typeface and Multiple Choice Option Format.
INSTITUTION	University of South Florida, Tampa.
PUB DATE	69
NOTE	9p.
EDRS PRICE	EDRS Price MF-$0.25 HC-$0.55
DESCRIPTORS	College Students, *Measurement Instruments, *Multiple Choice Tests, Perception, *Reading Comprehension, *Reading Research, Test Results, *Typewriting

ABSTRACT

The effects of typeface and item options arrangement on comprehension as indicated by multiple-choice test performance were investigated. Copies of the Ability to Interpret Reading Materials in the Social Studies, SRA Iowa Tests of Educational Development, Form X-4 were prepared in four typefaces: elite, pica, proportional, and script. For each typeface condition, the item options were presented either vertically (each option on a different line) or horizontally (each option following continuously across the page). Subjects were 80 college students randomly assigned to one of the eight treatment conditions. It was found that neither typeface nor item arrangement significantly influenced test performance. However, since the vertical item arrangement produced higher results for all typestyles, it was suggested that vertical vs. horizontal in combination with other format variables might significantly influence multiple-choice test performance, and further research on the physical arrangement of multiple-choice test items is recommended. Split-half and Kuder-Richardson reliability estimates are reported, and tables and references are included. (CM)

ED040820

TYPEFACE AND MULTIPLE CHOICE OPTION FORMAT

John Follman
Univ. of South Florida

A. J. Lowe
Univ. of South Florida

William Miller
Univ. of South Florida

Cronbach (1946) indicated that there are a number of response sets which influence a testee to obtain a different score from what he might obtain if the test items were presented in a different format. Most studies of such response sets have been concerned with personality characteristics and item difficulty arrangements. Relatively few reported studies have examined graphics, physical, test characteristics which might be lumped under the panoply of format, particularly typefaces, and also physical arrangement of items.

Payne (1967) reported that while there has been a substantial volume of research on typography, most of the research has been concerned with legibility, a smaller share with readability, and the remainder with specialized problems. He further noted that the reported research has compared typefaces commonly used in printing and not typefaces commonly used in typewriters. Poulton and Brown (1968) reported that apparently the only experiment comparing typefaces of teleprinters and typewriters was conducted by Fox (1963) who obtained significant differences for speed but not comprehension.

There has been considerable empirical examination of the effects of typeface variables on reading legibility. Tinker (1963) and Spache (1966) in reviews have indicated that an acceptable type format for legibility and reading speed is a 31 pica line with 2 point leading. Spache (1966) noted that in general type sizes between 9 and 12 points can be read equally fast and Tinker (1963) concluded that typefaces in common use are equally legible. The most commonly used typefaces in academic situations are elite and pica.

RE002 792

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

Insofar as type styles are concerned Bell (1939) found that cursive script was read significantly slower than manuscript. IBM (1967) reported that proportional spacing type, in which the characters are designed to four different widths thus allowing each character a tailored width, is read 6% faster than IBM Standard Prestige Elite.

Little has been reported on the effects of typeface on comprehension as indicated by multiple choice test performance, particularly in typical college and university settings. Poulton and Brown (1968) in England used two research designs, Separate Groups, and Greco-Latin Square to examine the effects of typeface on comprehension and reading rate. The Separate Groups Design examined: IBM 72 pica combined upper and lower case; IBM 72 pica lower case; Siemen's upper case; and IBM pica upper case. The Greco-Latin Square Design examined: Smith-Corona elite 66 upper and lower case combined; Siemen's upper case; Smith-Corona elite 66 upper case; and Dyeline Siemen's upper case. The dependent variables were 10 open-end comprehension questions for both designs, and rate, words per minute, for the Separate Groups Design. In the Separate Groups Design, the IBM 72 pica combined upper and lower case comprehension was 13% significantly higher than the average comprehension of the three upper case faces. There were small differences in comprehension between typefaces in the Greco-Latin Design although Smith-Corona elite 66 combined upper and lower case had the highest comprehension, slightly higher than Smith-Corona pica 1 lower case.

Payne (1967) in two experiments compared the effects of two kinds of typeface, proportional spacing (IBM Modern) and elite standard spacing (IBM Prestige Elite) on comprehension and speed of reading hard and easy passages of the Davis Reading Test. Standard spacing has all characters designed to one basic width while proportional spacing has each character designed to a tailored width. In most

comparisons across both experiments non-significant superior speed and comprehension were associated with proportional spacing. One comparison was significantly in favor of proportional spacing. Adult, non-student samples were used.

There has been little reported empirical examination of physical arrangement of items. Flaugher, et al. (1966) examined four arrangements of Scholastic Aptitude Test items likely to have answer sheet visibility. One finding was that some verbal item arrangements produced differences in difficulty levels for items not reached by some students. Spache (1966) cited Tinker's finding that arrangements of materials such as printing successive phrases on two lines, blocking the material, vertical or columnar arrangements, thought units horizontally spaced, all produced slower reading. Little empirical evidence of the effects of physical arrangement of test items is otherwise available.

This study had two main foci. One focus was on the effects of typeface on comprehension as indicated by multiple choice test performance. The second focus was on the effects of multiple choice item options' arrangements on comprehension. Specifically the objective was to determine the effects on performance in a multiple choice testing situation of four typefaces, ELITE, PICA, PROPORTIONAL, and SCRIPT each with item options presented either DOWN (each option on a different line for each item) or ACROSS (each option following sequentially, continuously for each item). It was anticipated that SCRIPT - DOWN and SCRIPT - ACROSS would be associated with inferior comprehension compared with the other six treatments. It was also anticipated that the ACROSS option format would be associated with performance inferior to that associated with the DOWN option format, for all four typefaces.

PROCEDURE

Subjects (Ss) were students in an introductory special education course at the University of South Florida in May, 1969. The Ss ranged from freshmen through

graduate students with median status being sophomore level. Ss were told that they were participating in a norming study of a social studies test.

Each S was randomly assigned to one of the eight treatment conditions. The Different Groups design was used as recommended by Poulton and Brown (1968) rather than a design with the same Ss receiving all treatments. The treatment conditions were: ELITE - ACROSS; ELITE - DOWN; PICA - ACROSS; PICA - DOWN; PROPORTIONAL - ACROSS; PROPORTIONAL - DOWN; SCRIPT - ACROSS; and SCRIPT - DOWN. The typefaces were: IBM PRESTIGE ELITE 72 Code 012, a 12 pitch weighted type; IBM COURIER PICA 72 Code 015, a 10 pitch square-serif pica type; IBM EXECUTIVE PROPORTIONAL, a type which allows each character a tailored width; and IBM SCRIPT Code 090, a 12 pitch type which simulates handwriting. The eight treatments were typed on stencils and reproduced on a conventional mimeograph machine adjusted for 20 pound paper and set for medium slow speed. Table 1 illustrates the ACROSS and DOWN multiple choice item options format arrangements using the sample item from Test 5 Ability to Interpret Reading Materials in the Social Studies, SRA IOWA TESTS OF EDUCATIONAL DEVELOPMENT, Grades 9-12, Form X-4 (Science Research Associates, 1960).

TABLE 1

ACROSS and DOWN Multiple Choice Item Options Arrangements

ACROSS	Which of the following is the largest island in the West Indies? A. Haiti, B. Jamaica, C. Cuba, D. Puerto Rico
DOWN	Which of the following is the largest island in the West Indies? A. Haiti B. Jamaica C. Cuba D. Puerto Rico

Test 5 Ability to Interpret Reading Materials in the Social Studies, SRA IOWA TESTS OF EDUCATIONAL DEVELOPMENT, Grades 9-12, Form X-4 (Test 5) (Science Research

Associates, 1960) was the experimental medium. The 53 item version of Test 5 was used as there was not enough time for all Ss to finish the 80 item version. The ITED Manual (Science Research Associates, 1963) reports a split half reliability estimate of .90 for the 53 item version of Test 5 for a twelfth grade sample.

A 4x2 ANOVA was conducted to determine if there were significant differences associated with typeface, option format, and interactions.

Split half and Kuder-Richardson Formula 20 reliability estimates were determined.

RESULTS

Table 2 indicates the means and N's for typeface, ACROSS vs. DOWN, and combinations.

TABLE 2

Means and N's for Typeface, ACROSS vs. DOWN and Combinations

	Elite	N	Pica	N	Proportional	N	Script	N	Total $\bar{X}$	N
ACROSS	32.30	10	32.81	11	34.13	8	35.45	11	33.68	40
DOWN	38.20	10	33.45	11	36.88	8	36.82	11	36.25	40
TOTAL $\bar{X}$	35.25	20	33.13	22	35.50	16	36.14	22		

Table 3 indicates the ANOVA for typeface, ACROSS vs. DOWN, and interaction.

TABLE 3

Analysis of Variance

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F
Typeface	3	109.96	36.65	.69
ACROSS vs. DOWN	1	132.61	132.61	2.51
Interaction	3	84.14	28.05	.53
Within Cells	72	3802.18	52.81	
Total	79	4128.89	52.26	

As Table 3 indicates, there were no significant F's. The relatively larger F computed for the ACROSS vs. DOWN difference, while not significant, reflects the difference in means between ACROSS and DOWN indicated in Table 2 which held in the ACROSS vs. DOWN comparisons for each of the four typefaces as well as when combined across the four typefaces. It appears that the arrangement of multiple choice item options may affect a testee's performance probably not substantially but conceivably enough to justify further investigation. This finding supports the speculation that other format factors may also influence multiple choice test performance. If this is the case it may be further speculated that multiple choice test performance may well be substantially influenced by format variables and that a considerable portion of multiple choice test variance may be spuriously associated with such format variables rather than with the desired knowledge and/or ability. While this situation is undesirable from a theoretical view, whether or not it is important practically depends on how the format variables interact with testees. If the format variables operate randomly the importance of format variables is minimal. However, if the format variables operate differentially with different testees, their test scores may be an inaccurate reflection of their real knowledge and/or ability. Further research to determine the effects of format variables in general and the differential effects of format variables in particular is suggested.

The highest typeface performance was associated with ELITE - DOWN, with ELITE - PROPORTIONAL, and ELITE - SCRIPT next. The effect of ELITE typeface washed out however when it was combined with ELITE - ACROSS.

The authors' hypothesis that SCRIPT would be associated with the lowest performance was not substantiated and in fact SCRIPT produced the highest performance in the ACROSS item options arrangement although none of these means differed much. The authors' second hypothesis that inferior performance would be associated with

the ACROSS item options arrangement was not supported statistically although the DOWN means were higher in every comparison.

The reliability of Test 5 was estimated so that if Test 5 was found to be reliable the experimental results could better be attributed to the independent variables. If Test 5 was found to be unreliable the authors would be unable to determine whether the lack of significant differences was due to the weakness of the independent variables or if there may have been effects which were masked by the unreliability of Test 5.

Since N's were small for the ACROSS and DOWN groups for each typeface and since there were no significant differences between ACROSS and DOWN groups, the ACROSS and DOWN Ss were combined into one group for the reliability estimates for each of the four typefaces.

Table 4 indicates odd-even split half and Kuder-Richardson reliability estimates, and N's for each typeface, with the ACROSS and DOWN groups combined.

TABLE 4

N's and Split Half and Kuder-Richardson Reliability Estimates
with ACROSS and DOWN Combined for Each Typeface

	N	Split Half	Kuder-Richardson
ELITE	20	.88	.87
PICA	22	.82	.84
PROPORTIONAL	16	.47	.56
SCRIPT	22	.91	.92

The high reliabilities associated with the ELITE, PICA, and SCRIPT typefaces reinforce the conclusion that insofar as comprehension is concerned the effects of commonly used typefaces are limited, small, that the typeface used probably does not substantially affect a testee's performance in a conventional multiple choice testing situation. This is consistent with the findings of Payne (1967) who obtained non-significant results in most comparisons and Poulton and Brown (1968) who

in some comparisons did get significant but small differences. The low reliability found for PROPORTIONAL may reflect the relatively small N which consequently might be non-representative, and also the small variance which obtained in the DOWN condition.

CONCLUSIONS

The following conclusions appear appropriate:

1. typeface does not substantially affect performance in conventional multiple choice testing.
2. while ACROSS vs. DOWN multiple choice item option format did not significantly influence test performance it was speculated that ACROSS vs. DOWN in combination with other format variables might significantly influence multiple choice test performance, and further research on the physical arrangement of multiple choice test items is recommended.

REFERENCES

Bell, H. M. The Comparative Legibility of Typewriting, Manuscript and Cursive Script: I. Easy Prose, Letters, and Syllables. Journal of Psychology, 8, 1939, 295-309.

Cronbach, L. J. Response Sets and Test Validity. Educational and Psychological Measurement, 6, 1946, 475-494.

Fox, J. G. A Comparison of Gothic Elite and Standard Elite Typefaces. Ergonomics, 6, 1963, 193-198.

Flaugher, R. L., Melton, R. S., and Myers, C. T. A Study of the Effects of Item Rearrangement. Research Bulletin, Educational Testing Service, Princeton, New Jersey, 1966.

International Business Machines Corporation. Once You've Signed the Letter, What You Say Is On Its Own. Form No. 543-0053, 4, 1967.

Payne, D. E. Readability of Typewritten Material: Proportional Versus Standard Spacing. The Journal of Typographic Research, 1, 1967, 125-136.

Poulton, E. C., and Brown, C. H. Rate of Comprehension of an Existing Teleprinter Output and of Possible Alternatives. Journal of Applied Psychology, 52, 1968, 16-21.

Science Research Associates. SRA IOWA TESTS OF EDUCATIONAL DEVELOPMENT, Grades 9-12, Form X-4, Test 5, Chicago, Illinois, 1960.

Science Research Associates. ITED, The Iowa Tests of Educational Development Manual for the School Administrator, Chicago, Illinois, 1963.

Spache, G. D. Good Reading for Poor Readers. Champaign, Illinois, Garrard Publishing Company, 1966.

Tinker, M. A. Legibility of Print. Ames, Iowa, Iowa State University Press, 1963.